DOCUMENT RESUME

ED 021 903 UD 004 462

By- Cloward, Robert D.
STUDIES IN TUTORING.
Columbia Univ., New York, N.Y. Social Work Research Center.
Pub Date 66
Note- 105p.
EDRS Price MF-$0.50 HC-$4.28
Descriptors- AFTER SCHOOL TUTORING, BEHAVIOR CHANGE, CHANGING ATTITUDES, CONTROL GROUPS, DATA, *DISADVANTAGED YOUTH, *ELEMENTARY SCHOOL STUDENTS, EXPERIMENTAL GROUPS, GRADES (SCHOLASTIC), *HIGH SCHOOL STUDENTS, INNER CITY, MASTER TEACHERS, PROGRAM EVALUATION, *READING ACHIEVEMENT, READING TESTS, *TUTORIAL PROGRAMS
Identifiers- Homework Helper Program, Iowa Silent Reading Tests, Mobilization for Youth, New York City, New York Tests of Growth in Reading

This is the report of a demonstration tutorial project which was conducted in New York City as part of the Mobilization for Youth program. Eleven tutorial centers were established in neighborhood elementary schools, and 240 students from local academic and vocational high schools were hired to tutor 544 fourth- and fifth-grade pupils. Each center was run by a master teacher, who, in addition to administrative activities, was responsible for training the tutors. For the purposes of evaluation tutors and tutees eligible for the program were randomly assigned to experimental and control groups. Experimental pupils were tutored either once or twice a week for 2 hours. Findings are reported concerning the impact of the program on the participants' reading achievement, school marks, behavior, and attitudes before and after their tutoring experience. Also presented are specific recommendations for the establishment of similar tutorial programs. Separate technical reports containing extensive evaluation data on both the pupils and their tutors are included. (LB)

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

STUDIES IN TUTORING

Robert D. Cloward

Research Associate

Research Center
Columbia University School of Social Work

1966

UD 004 462
ED021903

CONTENTS

Chapter I Introduction 1

Chapter II Mobilization's Tutorial Program 6

Chapter III The Impact of the Program 23

Chapter IV Guidelines for Tutorial Programs 36

Appendix A. Technical Report on the Pupils 48

Appendix B. Technical Report on the Tutors 72

CHAPTER I

INTRODUCTION

Although educators are not known for their unamimity in other respects, most would agree in assigning high priority to the problem of the elementary-school pupil whose severe reading retardation leads to progressive alienation from the school and ultimately from society. Of equal and perhaps even greater urgency is the problem of the older underachiever, the youngster whose lack of motivation to succeed in school results not only in low school grades but, inevitably, in a low-grade future. These problems are serious in any stratum, but they take on special significance when they occur among lower-class children, for in a free society, education is the principal means of achieving social and economic advancement. The purpose of this monograph is to describe one remarkably successful attempt to meet these problems and to suggest ways in which similar programs might be implemented in other communities.

* * *

In February 1963, Mobilization for Youth, Inc.,[1] in cooperation with the New York City Board of Education, instituted an after-school tutorial program for low-achieving pupils residing in a 67-block segment of Manhattan's Lower East Side. Once famed as a Jewish ghetto, the area in recent years has experienced an exodus of middle-class-oriented Jews and an influx of lower-class Puerto Rican and Negro families. The effects of these migrations are most noticable in the 16 public elementary schools in the area, where Puerto Rican youngsters in 1963 accounted for 59 percent of the pupil population and Negro youngsters for 17 percent.

By 1962, all of the elementary-schools in the area had been classified by the Board of Education as "special service" schools so as to permit reduced class size and expanded auxiliary services. The pupils in each school had the half-time services of a guidance counselor and a corrective reading teacher, plus increased access to other auxiliary staff.[2] However, even with the addition of these services, the reading achievement of pupils in the area remained far below that reported for pupils in the city as a whole. In 1963, 70 percent of the third-grade pupils in the neighborhood were reading below grade level as compared with 55 percent for New York City.[3]

Our research in the area indicated that the rates of reading retardation were fare greater for Puerto Rican and Negro pupils than for other youngsters. Among the third-graders, for example, 83 percent of the Puerto Rican pupils and 77 percent of the Negro pupils were reading below grade level, as compared with 51 percent of the white pupils. That this retardation was progressive as well as cumulative is evident from the eighth-grade statistics: 65 percent of the Puerto Rican pupils and 34 percent of the Negro pupils were reading three or more years below grade level as compared with 14 percent of the white pupils.[4] Clearly, a major effort was needed to arrest the reading retardation displayed by the vast majority of public-school children in the area.

The usual approach to the educational problem posed by low-achieving youngsters is to hire additional professional staff and build programs of remediation around them. This approach rests on the assumption that the problems of the retarded reader are so complex and intricate that only the professional reading specialist, by virtue of his training and experience, could be of real assistance. However, since a reading specialist can provide intensive service to only a few pupils at a time, it would be prohibitively expensive to hire as many of them as would be needed to solve the problem. Of perhaps greater importance is the fact that, the reading specialist, despite his superior technical skills, may be no better equipped to communicate with lower-class youth than is the middle-class-oriented teacher typically found in slum schools.

Mobilization's tutorial program rested on the assumption that older adolescents who come from the same social and economic strata as their pupils are in a better position than the middle-class-oriented teacher (or reading specialist) to understand and communicate with the children who live in our urban slums. Such youngsters, it was hypothesized, working under the direction of professional teachers, could serve as a bridge between the professional and the child in need, and could provide pupils with role-models who would enhance their aspirations for academic success.

Mobilization's approach, then, was to hire and train a large number of local high-school students to help low-achieving elementary-school pupils with their homework assignments and to tutor them in basic reading and other language-arts skills.

An integral aspect of the program was the systematic attempt to evaluate its effects on the pupils and on their tutors. The studies conducted between November 1963 and June 1964 provided substantive evidence that tenth- and eleventh-grade students, working as tutors under the supervision of expert teachers, not only produced significsnt gains in the reading skills of low-achieving fourth- and fifth-grade pupil but made still greater gains in their own reading achievement in a relatively short period of time. The results of these experiments have far-reaching implications for remedial education and for youth employment.

A detailed description of Mobilization's tutorial program is presented in the chapter that follows. Chapter III contains a brief description of the evaluation experiments, a review of their findings, and a discussion of their implications. Specific recommendations for the establishment of effective tutorial programs are offered in Chapter IV. The technical reports of the research have been included as appendices.

NOTES

1. Mobilization for Youth, Inc., was established in 1959 with funds from the National Institute of Mental Health, the Presidents' Committee on Juvenile Delinquency, the Ford Foundation, and the City of New York. It was a forerunner of the present Community Action Programs of the U.S. Office of Economic Opportunity.

2. Through funds provided by Mobilization for Youth, each school now receives the full-time services of a guidance counselor and a corrective reading teacher.

3. Citywide norms were obtained from Summary of Citywide Test Results for 1964-1965, Bureau of Educational Research, Board of Education of the City of New York, October, 1965, p. 9. The data on Mobilization-area pupils were obtained through a 1963 survey of the school records of 1,390 third grade pupils in 16 public schools.

4. Ethnic identifications were based primarily on family names, given names, and parents' place of birth.

CHAPTER II

MOBILIZATION'S TUTORIAL PROGRAM

The Homework Helpers Program, as Mobilization's tutorial project was called, was jointly funded as a demonstration project by the Ford Foundation and the City of New York; research and evaluation funds were provided by the National Institute of Mental Health.

The program was conducted on a pilot basis from February to June 1963. This period was used to solve organizational and logistic problems, to develop a curriculum of tutorial activities for the pupils, to identify research issues that should be explored, and to develop and field-test data-collection instruments. Nine tutorial centers were established in neighborhood elementary schools, and 110 eleventh- and twelfth-grade students from local academic and vocational high schools were hired to tutor 330 fifth-grade pupils one afternoon a week.

In the following year, when the main evaluation was conducted, the scope of the program was enlarged to include 190 tutors from the tenth and eleventh grades who were new to the program and 50 twelfth-graders who had participated in the program the preceding year. An entirely new population of pupils ---544 in all --- was drawn from the fourth and fifth grades of local elementary schools. Two additional centers were opened to accommodate the larger groups of tutors and pupils.

Structure of the Program

The tutoring took place from 3:15 to 5:15 on Tuesday, Wednesday, and Thursday afternoons. On Monday afternoons tutors participated in an in-service education program. The tutors were paid at the rate of $11.00 a week for six hours of tutoring and two hours of in-service training.

With only occasional exceptions, tutors in all centers worked on an individual basis with their assigned pupils. In five centers, each tutor worked with a different pupil on each of the three afternoons set aside for tutoring. In five other centers, only two pupils were assigned to each tutor, one reporting on Tuesdays and Thursdays and the other on Wednesdays. Each tutor-pupil pair was assigned to a specific classroom, with as many as four pairs to a room. After each session, the tutor escorted his pupil home.

Each center was administered by a "master teacher" recruited from the faculty of the school in which the center was located, from a nearby school, or from the District Office of the Superintendant of Schools. The master teacher was responsible for training the tutors, organizing and supervising tutorial activities, procuring and distributing educational materials, and coordinating the center activities with the activities and requirements of classroom teachers and other school officials. Applicants for this position were required to have at least five years' experience in teaching and to be free of conflicting after-school commitments. Most of the master teachers hired in September 1963 had served as center supervisors during the pilot-study period. All master teachers, veteran and novice met in October for a week of orientation and training conducted by the Homework Helper Program Coordinator. During the rest of the year, master teachers met monthly with the Program Coordinator to exchange information, plan in-service training activities for tutors, and discuss common problems.

Each master teacher was assisted by an attendance aide, a non-professional member of the community, often a parent from the local school. The attendance aide welcomed tutors and pupils as they reported to the center, kept attendance records, helped to distribute and collect educational materials, and in other ways assisted the master teacher. The attendance aides were required to have completed at least the sixth grade and to be able to read and write well enough to carry out their clerical duties.

Recruiting and Selecting the Tutors

The tutors were recruited through an extensive publicity campaign, which included radio announcements, newspaper articles, posters, and flyers distributed to tenth- and eleventh-grade students in local and nearby academic and vocational schools. Most of the tutor applicants were residents of the project area who were attending the local academic high school. Applicants who were reported by the schools as being emotionally unstable or whose parents would not consent to their serving as tutors were eliminated from consideration.

In the early discussions about the program, shcool officials had advocated the establishment of relatively high academic requirements for tutors, in the belief that only youngsters' who were themselves successful in school could serve effectively as inspirational models for low-achieving pupils. Mobilization, however, took the position that low-achieving high-school students might be better able to understand and empathize with low-achieving elementary-school pupils than would high-school students to whom learning comes quickly and easily. Furthermore, it was clear that defining tutor eligibility in terms of better-than-average school achievement --- or even reading achievement at or above grade level --- would eliminate from further consideration most of the Negro and Puerto Rican applicants. In our view, the main vehicle for change would be the pupil's identification with his tutor as a role-model, an identification that would be greatly facilitated by ethnic pairing of pupils and tutors. Since the majority of pupils needing service were Negro or Puerto Rican, it was believed necessary to hire as many Puerto Rican and Negro youngsters as tutors as possible. The eligibility requirements, therefore, were set at satisfactory (passing) marks in school and reading achievement no lower than three years below grade level when last tested by the schools. Nevertheless, of the 240 youngsters accepted for tutorial positions in 1963, only 19 percent were Puerto Rican and 18 percent were Negro, 2 percent were Oriental. The remaining 61 percent were white. Seventy percent of the tutors were girls; 43 percent were tenth-grade students, 36 percent were eleventh-grade students, and 21 percent were twelfth-grade students.

Selecting the Pupils

The pupils who received tutorial service came from a population of 2500 fourth- and fifth-graders in project-area schools who were reading below grade level when last tested by the schools.[1] A letter in both Spanish and English was sent to the parents of these youngsters explaining the tutorial program and inviting them to apply for their children. The guidance counselors in each school were asked to screen the applicants to eliminate pupils with very serious behavioral problems, pupils classified by the school as mentally retarded, and pupils with long histories of truancy.

Over the year, 544 pupils were accepted for tutoring. Their ethni distribution was 60 percent Puerto Rican, 28 percent Negro, 9 percent white, and 3 percent Oriental. Fifty-four percent of the pupils were boys, as compared to 30 percent of the tutors. Fifty-six percent were fourth-grade pupils.

Training the Tutors

Special emphasis was placed on training the inexperienced high-school students, before and during the program, to serve effectively as tutors. The general goals of the tutor training program were as follows:

1. To orient tutors to the purposes of the program and the responsibilities of a tutor.
2. To help beginning tutors to overcome their initial anxiety about serving as tutors.
3. To help tutors to establish good working relationships with their pupils and to treat their pupils with warmth and acceptance.
4. To train tutors to study their pupils systematically and objectively.
5. To provide tutors with the information they need for effective service.
6. To give the tutors vicarious experience with a variety of tutorial situations and problems.

Pre-service Training

Before being introduced to their pupils, the tutors were given 16 hours of training in tutorial skills (two hours a day, four afternoons a week, for two weeks). All the tutors and their master teachers met together with the Program Coordinator during the first week for lectures and panel discussions. During the second week, tutors met with their master teachers in the centers to which they had been assigned.

The pre-service training experiences focused on the goals of the tutorial program, the organization of the program, the duties of a tutor, the characteristics of the pupils (including the scholastic levels at which they probably would be operating), and the kinds of activity in which tutors might engage their pupils in the first few tutorial sessions.

Many of the tutors had expressed concern about their ability to give effective service to children who were experiencing considerable difficulty in their school work. Recognizing that much of what tutors would need to know could be taught to them in later sessions, master teachers focused pre-service training on the routine procedures and activities that should be followed in the first few sessions with the pupils. By concentrating on things the tutors could do in their first meetings with their pupils, the master teachers helped greatly to allay the tutors' early concern.

In-service Training

In the weekly in-service training meetings held during the rest of the year, the master teachers attempted to familiarize the tutors with the elementary-school curriculum so that they could help pupils with homework assignments; to acquaint them with specific techniques and materials useful in teaching reading and language-arts skills; and to help them deal effectively with problems arising out of their relationships with pupils, to interpret pupil behavior, and to recognize and respond to pupil needs.

Early training sessions tended to focus principally on the development of good tutor-pupil relationships, the systematic study of the pupil, and the mechanics of tutoring. Tutors were encouraged to get to know their pupils as individuals with specific needs and interests and to use each pupil's needs and interests as vehicles for learning. They were cautioned not to exhibit dismay at how little their pupils seemed to know but to accept the pupils' initial reading level and select interesting activities that would move them toward improved performance. Tutors were warned to avoid using coercion, threats, or rebuke to induce recalcitrant youngsters to do their homework or engage in other scholastic activities.

A good deal of time in these early training sessions was devoted to a review of the elementary curriculum. To be of assistance to their pupils, tutors needed to know what was being taught at each grade level in the schools. Particular care was taken to see that tutors did not use methods or activities that would conflict with what was being taught in the classroom.

Mathematics is a case in point. Teachers in the area were using a developmental system of mathematics instruction that was entirely foreign to the tutors. Trained in traditional arithmetic reasoning, the tutors could do little more than check arithmetic homework assignments for correct answers. Thus, before tutors could meaningfully assist pupils who were having difficulty completing mathematics assignments, several in-service meetings had to be devoted to training the tutors in the thought processes involved in developmental mathematics.

Similarly, a good deal of time was devoted to teaching the tutors methods of reading instruction. Tutors were briefed on phonic skills useful in analyzing new or unfamiliar words. They were shown how to make and use flash cards to improve sight recognition of whole words as well as recognition of initial consonants, consonant blends, word endings, and vowel sounds. They were taught when and how to use commercially prepared reading materials and games. Although the master teachers made many suggestions, the emphasis was always placed on encouraging the tutors to use their own creative talents to devise interesting materials and activities for reading instruction and practice.

Throughout the year, tutors were asked to write anecdotal or narrative reports of their experiences with their pupils. These reports proved to be invaluable to the master teachers. Not only did they provide information about the developing relationships between tutors and pupils, but they often helped to pinpoint common problems that needed to be discussed in the in-service training sessions. Here are a few examples of anecdotal reports early in the year:

November 27th:

Manuel came late and explained that his mother wanted him to go home and change his clothes before he came here. I was working with Victor, whose tutor was absent. Manuel didn't have any homework, so I gave him and Victor a spelling bee.

Manuel kept calling out the words even when it wasn't his turn. At one point he misspelled a word. When I wouldn't let him try again, he got angry and refused to spell the next word. I completely ignored him until he came over and apologized.

Then the three of us played a vowel game until it was time to go home.

December 4th:

Pamela did not have any homework, so I started teaching her phonics. She knew almost all the sounds.

I asked her to tell me what time it was from the clock on the wall. Her answer was wrong by twenty minutes. So I started teaching her how to tell time.

Soon she got restless, so I gave her a spelling test consisting of ten words. She got nine of them right.

On the way home, we talked about whether she understood what I had been teaching her so far. She said she understood almost everything except how to tell time.

December 10th:

George had homework in math. He had some difficulty with subtraction. It took 30 minutes for him to finish his homework. Then I had him read from the Reader's Digest /skill text/. He answered all the questions correctly and said he liked the story.

George asked me if we could play a game. We played Scrabble until it was time to go home. He seems to like this game very much.

The master teachers used anecdotal records as illustrations of good and bad tutorial practices, as situational case studies to provide vicarious experience in tutorial roles, and as a basis for group discussion. Tutors projected their own thoughts and feelings into the situation described and compared their perceptions and judgments with those offered by other tutors.

The in-service education experiences were often supplemented by individual conferences with the master teacher. When a master teacher perceived that a tutor was having difficulty in developing a warm, accepting relationship with his pupil or controlling his pupil's behavior, the teacher arranged a conference with the tutor.

Most of the master teachers made it a practice to hold periodic individual conferences with their tutors. These conferences provided opportunities for master teachers to give tutors some guidance with their personal problems as well as the problems presented by their pupils.

On one Monday afternoon each month, the tutors from all the centers met together with all the master teachers for a special program. Some of these meetings involved lectures or panel presentations by professional educators. Occasionally tutors were given an opportunity to demonstrate special materials that they had developed and to discuss ways in which they could be used effectively.

The Tutorial Sessions

As we have noted, the tutors usually spent part of each session helping the pupils with their homework assignments. These assignments consisted principally of spelling drill and arithmetic problems, with occasional assignments in reading, social studies, and science. Emphasis was placed on helping the pupil to help himself by developing good work habits and study skills.

The tutors' main objective, however, was to improve their pupils' language skills. To this end, they used a variety of methods, changing their techniques and materials to suit the pupils' changing needs. The first few sessions were used largely to establish rapport with pupils, to lay a basis for future working relationships. In these meetings, pupils and tutors played games together or talked. Soon pupils began to bring some of their homework assignments to the sessions, seeking their tutor's assistance. As the pupils became more confident and skilled, they began to do their homework themselves, consulting the tutors only if they were having difficulty. Reading and creative activities were gradually added to the tutoring sessions. By the end of the second month the typical tutorial session consisted of 30 minutes spent on homework, 30 minutes on reading, 15 to 30 minutes on games and recreation, and 15 minutes on refreshments, roll-taking, and other non-tutorial activities. This pattern was sustained throughout the remainder of the year.

Two sets of graduated exercises were used to provide practice in reading. Science Research Associates Reading Laboratory #2 consists of a series of very finely graded reading selections followed by questions which test the pupil's comprehension of the passages. As the pupil achieved competence at one level he was moved to a slightly more difficult level. The second set of reading exercies, the Reader's Digest Skill Texts, is similar to the SRA Reading Laboratory except that the selections are not so finely graded in difficulty. Additional reading practice was provided by Viewlex filmstrips with captions, by library books, and occasionally by social-studies assignments.

Attempts were made to develop the pupils' language skills through creative writing exercises. A tutor might have his pupil select a picture from a file and write a story about it or draw a picture and write an appropriate sentence or two underneath it. A few pupils were encouraged to write plays for puppet shows and to record them on tape. Often pupils wrote stories on topics suggested by their tutors, such as "How I Will Spend My Summer Vacation," "My Favorite Game," "My Pet."

These creative writing activities served not only to stimulate self-expression but also as a means of individualizing instruction for each pupil. Tutors working with severly handicapped youngsters sometimes began by having the pupil dictate a story consisting of a few sentences and then read the story aloud as the tutor had recorded it. More advanced pupils began by writing their own brief stories and then, with the help of the tutor, editing what they had written for spelling, sentence structure, and punctuation. As pupils began to write articles and stories of two or more paragraphs, tutors found opportunities to explain paragraph organization, unity and coherence.

Newspapers were published in each center as motivational devices. Both pupils and tutors were encouraged to contribute. Indeed, much of the creative-writing activity was focused on producing articles, stories, and poems for the newspaper. Pupils obviously derived great satisfaction from having the fruits of their efforts displayed publicly in the newspaper. Many of them carried copies of the paper around with them to show their friends.

Considerable emphasis was placed on teaching phonic skills as an aid to breaking the code for new and unfamiliar words. Tutors made flash cards and word lists illustrating word families by initial consonants, consonant blends, vowel sounds, and word endings. Pupils were encouraged to keep notebooks in which they recorded new and unfamiliar words.

Extensive use was made of such educational games as Scrabble, Lotto, and Rolling Readers. These not only provided practice in vocabulary, spelling, sentence structure, and other language skills but made the tutorial sessions more enjoyable.

Costs of the Program

During the 1963-64 academic year, the total amount spent on the Homework Helper Program was $141,688.00. As is usually the case with educational programs, most of these funds (85 percent) were used to pay the salaries of the professional and non-professional staff members employed in the program. Tutor salaries alone accounted for $61,753.00, or almost 44 percent of the total expenditure. Master teachers, who were paid at the rate of $6.00 per hour, accounted for 17 percent of the total. Attendance aides, who were paid $1.75 per hour, accounted for only 2 percent. (See Table 1.)

The economy-minded educator may be quick to point out that the program might be replicated at half the price by using volunteer tutors. In our view, the demands placed on the tutor go far beyond what might reasonably be expected from unpaid volunteers. Indeed, we have since raised tutor salaries to provide increased financial rewards for more experienced tutors.[2]

Based on our expenditures and the number of pupils served, our best estimate of the costs per pupil per hour of instruction is $4.13.[3] This figure is far below what would have been incurred had the equivalent amount of individual assistance been given by professionally trained reading teachers.

TABLE 1

Program Expenses
1963-64

Staff	
Salary program coordinator	$10,539.00
Clerical services	10,000.00
Salaries, 12 master teachers	23,878.00
Salaries, 11 attendance aides	3,000.00
Salaries, 240 student tutors	61,753.00
Staff fringe benefits	1,305.00
Insurance for staff, tutors, and pupils	6,038.00
Social Security and taxes	3,937.00
	120,450.00
Overhead and Operating Expenses	
Rent, M.F.Y. office space	1,283.00
School custodial services	7,355.00
Light, power	172.00
Office supplies and equipment	3,014.00
Telephone, telegraph	533.00
Postage	200.00
Travel expenses	32.00
Books & periodicals	208.00
Educational equipment & supplies	7,426.00
Pupil refreshments	1,015.00
	21,238.00
TOTAL	$141,688.00

NOTES

1. These scores often were based on tests administered in the third-grade and did not necessarily reflect the current status of the pupils. Of those accepted for service in the program, approximately 10 percent subsequently were found to be reading at or slightly above grade level.

2. The original rate was $1.50 per hour for tutorial activities and $1.00 per hour for training. The pay scale has since been raised to $1.50 per hour for beginning tutors regardless of whether the time is spent in tutoring or training; $1.75 per hour for tutors with one years' experience; and $2.00 per hour for tutors with two year's experience.

3. This figure was derived by dividing the total year's expenditure of $141,688.00 by 33,120, the number of pupil hours of instruction given in the program. This latter figure was obtained by multiplying the total number of hours of instruction given that year (160) by the average daily attendance (207).

CHAPTER III

THE IMPACT OF THE PROGRAM

During the period of planning for the tutorial program, a number of questions were raised concerning the effect the program might have on the tutors and their pupils. In view of the complex nature of reading disabilities, some school officials anticipated that tutoring by high-school students would have no effect on the reading skills of low-achieving elementary-school pupils. Furthermore, to the degree that the high-school student became engrossed in his tutorial activities, they believed that he was likely to neglect his own school work, with disastrous results. In this view, improvement in pupil achievement, if it occurred at all, would be obtained at the expense of the tutor's school achievement.

In order to test these assumptions empirically, two concurrent experiments were conducted between November 1963 and June 1964. One experiment attempted to assess the impact of the program on the pupils. The second was concerned with determining what effect the tutorial experience had on participating high-school students.

Both studies were structured as classical experiments, with random assignment of eligible subjects to experimental and control situations. Experimental pupils were provided with tutorial services on a one-day or two-day-a-week basis. Control pupils were systematically denied service in the program. Similarly, high-school students in the experimental group were hired as tutors, and control high-school students were told that they could not be given jobs as tutors. In both studies, the control subjects served as a standard against which the achievement of the experimental subjects, before and after service, could be evaluated. The legitimacy of these comparisons depends in no small measure on the fact of random assignment. This process not only precluded the operation of bias in the selection of samples but served as well to ensure the comparability of experimental and control subjects, for the distribution of all known and unknown factors was left entirely to chance. In the absence of initial differences between experimentals and controls, differential outcomes may legitimately be attributed to differential treatment.

In out view, programs which are intended to improve the educational status of participants should be evaluated on the basis of objective measures of performance rather than --- as is often the case --- on the basis of such tangential data as program-attendance statistics, unsystematic observations of individuals, or the subjective judgments of participants. The present studies therefore focused on the "hard facts" of reading achievement. Although attempts were made to measure objectively changes in the attitudes, aspirations, and values of the participants, these data were viewed as secondary outcomes which could be accepted with confidence only if they correlated with reading achievement.

The Pupil Study[1]

The findings on the impact of the program on the pupils were based on the study of 356 experimental and 157 control subjects drawn from the population of fourth- and fifth-grade pupils deemed eligible for service. Included in these samples were 100 subjects who were to receive tutorial services two afternoons a week (the four-hour treatment group), 73 subjects who were to be tutored only one afternoon a week (the two-hour treatment group), and 79 control subjects, all randomly drawn from the same schools. The remaining 183 experimental subjects were tutored only two hours a week. They were compared with a group of 78 controls who received no tutoring.

Instrumentation

Pupil reading achievement was measured on a before-and-after basis, using the New York Tests of Growth in Reading, Level G, Form 1 (revised).[2]

To determine the effects of the program on school marks and school behavior, data were collected on teacher assigned grades in reading, spelling, mathematics, science, and social studies; on behavioral ratings in work and study habits, social behavior, and written and oral participation; and on school attendance. This information was gathered from school records for the 1962-63 (prestudy) and 1963-64 (post-study) academic years.

Before-and-after data on pupil attitudes and aspirations were gathered through the administration of a 44-item questionnaire. Since many of the pupils could not read well enough to take the questionnaire by themselves, the entire instrument was read to all subjects.

Two aspects of the tutor-pupil relationship were investigated: the possible effect of sex and/or ethnic matching of pupils and tutors on the reading achievement of the pupils, and the possible effect of the pupil's attitudes and feelings toward his tutor. The latter was measured by means of a research schedule containing 42 questions, so constructed as to permit the expression of negative as well as positive feelings and attitudes.

The Findings

At the beginning of the program, the average fourth-grade subject was reading at grade 3.5, eight months below grade level, and the average fifth-grade pupil was reading at grade 4.2, one year and one month below grade level. There were no significant differences between experimental and control subjects.

After five months of tutorial instruction, pupils in the four-hour treatment group showed significantly greater improvement that did control subjects. In terms of grade equivalents, the four-hour pupils showed an average of six months' reading improvement in five months' time whereas the control pupils showed only three and a half months' growth during the same period. The growth rate of the controls -- or, more accurately, their progressive retardation -- approximates the average rate for Puerto Rican elementary-school children in the Mobilization area. Thus the four-hour group not only arrested their retardation but began to catch up. The comparable two-hour group made a gain of five months in reading during the same five-month period. Although this growth rate exceeded that of the controls, the difference was not statistically significant.

The remaining 183 experimental subjects, all of whom were in centers offering only a two-hour exposure to tutoring, showed only 3.8 months' growth as compared with 3.9 months for their respective controls.

Further analysis of the reading data indicated that the four-hour treatment was most effective for those pupils who were the most severely retarded at the beginning of the study. The degree of change was not associated with such factors as the pupil's sex, grade in school (fourth or fifth), or access to school programs in reading remediation. Although ethnicity was not in general associated with improvement, Negro pupils who were matched to their tutors in sex and ethnicity showed significantly greater gains in reading than did control Negroes.

No significant before-and-after differences were found in pupil school marks, attitudes toward school and school-related activities, educational aspirations, or social values, even when the degree of reading improvement was taken into consideration. Nor was any significant relationship found between pupil attitudes toward their tutors and pupil reading achievement.

The Tutor Study[3]

The tutor study utilized random samples of 97 experimental subjects and 57 control subjects,[4] both drawn from the eligible population of tenth- and eleventh-grade applicants for tutorial jobs. None of the subjects had ever served as tutors before.

Instrumentation

To estimate the effects of the program on the tutors, before-and-after-program data were collected on reading skills, school achievement, attitudes, values, educational aspirations, and interest in becoming a teacher.

The Advanced Level of the Iowa Silent Reading Tests (Revised New Edition)[5] was used to measure changes in the reading skills of the subjects. Although this reading battery is rather long and exceptionally difficult to administer, it is especially valuable in that it yields eight test scores covering a wide variety of reading and study skills.

School-achievement data were obtained from school records for the 1962-63 academic year (prestudy) and the 1963-64 academic year (poststudy). Data on attitudes, values, aspirations, and interest in teaching were obtained through the administration of a 54-item questionnaire developed by the present writer.

The Findings

Much to our surprise, the program had a major effect on the reading achievement of the tutors. As measured by the Iowa Silent Reading Tests, the experimental subjects had a prestudy mean reading level of grade 9.9 as compared with a mean level of grade 10.1 for the controls. Since the mean grade of origin for both samples was 10.7 the average experimental subject was reading eight months below grade level and the average control subject was reading six months below grade level. However, both samples showed considerable range in reading achievement. Twenty percent of the experimentals and 21 percent of the controls were reading below the eighth-grade level.

In the seven months between administrations of the research instrument, the tutors showed an average gain in reading of three years, four months, as compared with one year, seven months, for the control subjects. The remarkable growth of the controls probably was an artifact of the testing situation, for the complexity of the directions for taking the test probably had the effect of artificially depressing the prestudy scores for all subjects. Since an alternate form of the same test was used in the poststudy administration, it may be that a substantial portion of the increase for both groups was due to their increased familiarity with the test directions rather than with the test items. This factor, however, does not explain away the sizable differential between the growth rates of experimentals and controls. Even if we take the control mean as an estimate of the effects of normal growth plus increased familiarity with the test situation, the net effect attributable to the experimental treatment is a reading increment of one year, seven months.

The data for the tutors, as for the pupils, clearly indicate that the effects of the experimental treatment were maximized for subjects with initially low reading skill. The effects of verbal aptitude,[6] grade of origin, sex, and ethnicity were not found to be significant.

No significant differences were found between tutors and their controls on before-and after-program measures of school marks, attitudes toward school and school-related activities, social values, educational aspirations, or interest in becoming a teacher. Nor did the experimental subjects who showed high change in reading.

The failure of the tutorial experience to influence the school marks of the tutors is particularly interesting. The reader will recognize that our poststudy measure of school achievement consisted of school marks earned while the subjects were serving as tutors. The measure therefore focuses on concurrent rather than post-treatment effects. Although we cannot be certain, it is reasonable to assume that the high reading gains made by the tutors will enable them to earn higher marks in their future school work. At any rate, it is clear that service as a tutor did not adversely affect school achievement.

The Prediction Studies

Two substudies were conducted in an attempt to establish guidelines for the selection of future tutors.

The first substudy was an attempt to define the effective tutor by exploring the possibility of predicting pupil gain in reading given a knowledge of the prestudy characteristics of their tutors. The second substudy was concerned with the issue of selecting as tutors youngsters who will derive maximum benefits from the tutorial experience.

To answer the first question, the intellectual, attitudinal and demographic characteristics of 38 tutors were used to predict the reading gain of their four-hour pupils. The intellectual characteristics included the tutor's prestudy reading level, verbal aptitude, and prestudy marks in school. The demographic variables consisted of the tutor's sex, ethnicity, and grade in school. The attitudinal variables consisted of the tutors' prestudy response to items in our research questionnaire concerned with attitudes toward school, interest in reading for recreation, interest in becoming a teacher, and educational aspirations.

None of these factors, singly or in combination, was found to correlate with the reading gain of the pupils. Although our sample for this study was rather small, the data lead us to conclude that pupil reading gain was not related to or influenced by the demographic, intellectual, or attitudinal characteristics of their tutors. Pupils working with very bright, high-achieving tutors made no greater gains than pupils whose tutors lacked high intellectual qualifications. Eleventh-grade tutors were no more effective than tenth-grade tutors. Likewise, tutors with high positive attitudes toward school and high educational aspirations were no more effective than tutors with less positive attitudes and lower aspirations.

The second substudy was an attempt to discover what characteristics our high-change tutors had in common. The study utilized a sample of 71 experimental tutors and examined the same intellectual, attitudinal, and demographic variables as the first study.

No single factor showed high predictive value, but all the factors combined predicted tutor reading gains far beyond the level of chance (multiple r = .52). The two best predictors were pre-study reading level and verbal aptitude (multiple r = .41). The addition of other factors increased the accuracy of prediction only slightly. These data indicated that the tutors who made gains in their own reading skills tended to have high verbal facility plus low initial reading skills.

Implications of the Studies

It is clear from the findings of these studies that to be effective, tutors do not need twelve years of formal education and extensive training in reading pedagogy. Nor need they be highly successful in their own school work. The average high-school student can learn to be an effective tutor.

Contrary to expectations, high-school tutors are most effective with those pupils who are the most severely retarded in reading. These are the youngsters who, because of their unsatisfactory progress in school, have come to expect ridicule, rejection, and continued failure. Teachers tend to regard these children as a burden and are reluctant to spend class time in an attempt to teach them the basic skills that they failed to learn in earlier grades. In a tutorial situation, where emphasis is placed on individual attention and basic skill training, these youngsters can make remarkable progress in reading.

Clearly, the major impact of the tutorial experience was on the tutors themselves. This finding has enormous implications both for education and for youth employment. Tutorial programs not only can provide older youth in a low - income area with gainful employment but can serve to upgrade their academic skills as well. Indeed, the astonishingly high reading gains made by tutors who were reading far below grade level at the beginning of the study raise the intriguing question of whether high-school dropouts might be successfully employed as tutors, not just to help under-achieving elementary-school pupils, but to improve their own academic skills. Attempts to remedy the dropouts' educational deficiencies by placing them in pre-employment training programs have not been notably successful. Having experienced failure and humiliation in the classroom and being alienated from school, these youngsters tend to rebel against learning situations in which they are cast in the role of a student. Assigning tutorial roles to such adolescents might help to make learning enjoyable and profitable for them. Obviously, youngsters who are functionally illiterate cannot be employed as tutors. However, many early-school-leavers read well enough to perform simple tutorial tasks under expert supervision. To the degree that serving as a tutor helps these youngsters improve their own academic skills, many could be rescued from the certainty of a bleak future.

NOTES

1. Technical data for the pupil study are presented in Appendix A.

2. Bureau of Educational Research, Board of Education of the City of New York, 1959.

3. Technical data for the tutor study are presented in Appendix B.

4. Control subjects were paid five dollars for their participation in the testing program.

5. Harcourt, Brace, and World, Inc., 1939.

6. An experimental edition of the Borgatta-Corsini Quick Work Test (Harcourt, Brace and World, Inc., 1962) was administered at the beginning of the program to measure the verbal aptitude of experimental and control subjects. The test was not re-administered at the end of the study.

Chapter IV

Guidelines for Tutorial Programs

In the previous chapters, we have focused on Mobilization's tutorial program, its methods, and the studies which proved its effectiveness. In this chapter, we offer specific recommendations for those who wish to set up similar programs. In each case, we have tried to explain our reasons for the recommendation on the basis of empirical evidence and three years' experience with the Homework Helpers Program. It should be made clear, however, that in very few instances have we tested alternative methods. What we offer here is a list of the techniques and arrangements that Mobilization has found to be effective.

There are no rubrics in this chapter; each numbered recommendation stands by itself. However, for the reader who is interested in specific phases of tutorial programs, we offer the following rough classification: items 1-8 deal with the selection, training, and deployment of tutors; items 9-12 focus on the selection of pupils and their assignment to tutors; items 13-15 are concerned with the duties and qualifications of supervising teachers and attendance aides and items 16-18 discuss the physical setting of the tutorial activities. Item 19 is concerned with evaluation.

1. Whether the major objective of the program is improvement in the tutors or in the pupils, tutors should be given genuine responsibilities for helping the pupils. In Mobilization's experience, the dramatic improvement in the tutors' reading levels were achieved indirectly, as a result of the tutors' attempts to help their pupils. There is no reason to believe that hiring teenagers merely to supervise or "baby-sit" with elementary-school pupils will have any effect on the tutors -- or on the pupils either. On the other hand, there is good reason to believe that when teenagers are charged with the task of trying to reach and help a child who is having difficulty, they not only can learn to give subtantive assistance to the child but will be helping themselves as well.

This is not to say that the potentials in a tutorial situation for helping tutors should be ignored or even that they should be subordinated to the goal of improving pupil achievement. Indeed, the real goal of the program might be to improve the educational achievement of the tutors. We are suggesting that this can be accomplished only by giving the teenagers a full measure of tutorial responsibility.

2. Tutors should work with pupils on an individual basis. Since the present study focused on a tutorial situation in which tutors and pupils met on a one-to-one basis, we have no evidence that having each tutor work with more than one child at a time

would be effective. On the other hand, the individual attention characteristic of a one-to-one relationship appears to have a very positive effect on pupil achievement.

3. Each pupil should receive at least four hours of tutoring a week. This recommendation follows logically from the major finding in our study of the pupils. Two hours of tutoring weekly is not sufficient exposure to produce change in achievement within a reasonable period of time.

4. Tutors should be selected primarily on the basis of their interest in serving as tutors. In Mobilization's program, tutors with high intellectual and academic qualifications were no more effective in helping pupils learn to read better than were tutors with relatively low qualifications. Preference, however, should be given to applicants with lower than average reading achievement, for these are the youngsters who are most likely to benefit from their experience as tutors.

5. Considerable emphasis should be placed on the training of the tutors. Few applicants for tutorial positions will possess at the outset all of the information, skills, and understandings they will need to deal effectively with low-achieving children. The employment of lay tutors presupposes a pre-service training program and a series of on-going in-service training experiences.

6. Attention should be given to the selection of pedagogical methods that are appropriate to the goals and content of the tutor training experiences. "Telling" is not always the best method of teaching. While lectures are a useful way of imparting knowledge, they have little value in developing tutorial skills or altering the ways in which tutors view their pupils. For these purposes, having tutors role-play typical tutorial situations might prove far more effective. Since learning at any level requires the dynamic cooperation of the learner, emphasis should be placed on methods which encourage active participation on a meaninful level.

7. If possible, the tutors should be paid for their services. We have noted that a minimum of four hours of tutoring a week appears to be necessary for pupil improvement. Thus a commitment of at least six hours a week (including two hours of in-service training) is required of the tutors, not to mention a period of pre-service education. Programs which depend on unpaid volunteers are likely to find many tutors dropping out to take part-time jobs, for recreational pursuits, or to meet other commitments. Paying the tutors not only meets the need of many low-income youngsters to find gainful employment while in school, but also provides them with material recognition for the important work they are doing.

8. If tutoring is to take place in the late afternoon or early evening, tutors should be required to escort their pupils home. This practice not only ensures the safe return of the pupil to his home, but offers an opportunity for tutor and pupil to get to know each other on a less formal basis. It also enables the pupil's parents to meet the tutor and thus provides a link between parent and center.

9. The recipients of the tutorial service should be selected on the basis of need, with preference given to those who are the most severely retarded in basic communication skills. We noted earlier that pupil reading improvement in Mobilization's tutorial program tended to be associated with initial reading level. Of the four-hour pupils, those initially reading at grade 3.3 and below showed an average of seven months' gain in reading as compared with only five months' gain for those reading at grade 3.4 and above. These data suggest that our high-school tutors were more effective when working with beginning readers than with advanced readers.

10. All pupils assigned to a given tutor should be reading at the same level. If a tutor is to work with several pupils each week, the pupils should be assigned on the basis of functional reading level. The tutorial activities for pupils reading at the fourth- or fifth-grade level differ markedly from those appropriate for children reading at the second-grade level. Tutors are likely to be more effective when all the pupils assigned to then are

operating at the same grade level.

11. In general, it is desirable to assign male tutors to boys. The studies found that assigning Negro pupils to Negro tutors of the same sex tends to maximize the achievement of the pupils. For other ethnic groups, sex-ethnic matching of pupils and tutors does not seem to affect achievement. However, many of the boys in our pupil sample, regardless of ethnicity, objected to being assigned to a female tutor, apparently feeling that their status as males was being threatened. Most of these youngsters got over these feelings once they got to know their tutors; in only a few instances was it necessary to reassign them to male tutors.

12. The supervising teacher must be prepared to give counseling when pupil-tutor difficulties arise and to make reassignments if necessary. Supervising teachers often can make excellent matches on the basis of their assessment of the personalities of pupils and tutors. Information provided by classroom teachers can be invaluable for this purpose. Once the matching has been made, the master teacher should closely supervise the developing relationship between tutor and pupil. If conflict appears, the teacher should interview the participants to determine whether counseling the tutor will resolve the difficulty or whether reassignment is necessary. If the teacher concludes that the pupil must be directed to another tutor, this should be done as early in the program as possible.

13. The supervising teacher must be on hand at all times, both to help and advise student-tutors as needed and to gather observational material for use in in-service training. Lay tutors, particularly youngsters, cannot be expected to know how best to handle the educational and disciplinary problems that inevitably arise. A responsible adult must be available to whom tutors can turn for advice and guidance. Furthermore, if the in-service education provided for the tutors is to be effective, the person responsible for the program must have an opportunity to observe each tutor in action, so that the trainer can tailor the in-service experiences to the needs and interests of the tutors.

14. Supervising teachers must be carefully chosen, on the basis of personal as well as professional characteristics, and must be adequately compensated. The personal and professional qualities of the supervising teacher have much to do with the effectiveness of the tutorial services rendered in the center. The supervising teacher not only must be qualified to diagnose the learning dysfunctions of each pupil and to develop effective remedial activities, but also must be expert in interpreting pupil and tutor behavior and responding to their needs. The teacher must himself be secure enough to be able to give the tutors status as semi-professionals in recognition for the important work they do.

The teacher's most important task is to train inexperienced tutors to play helping roles for their pupils. He must therefore be able to organize his administrative duties and delegate responsibilities to attendance aides so as to concentrate on the important work of training and supervising tutors.

There are advantages to recruiting supervising teachers from the faculties of local schools. Not only do such teachers have intimate familiarity with the curriculum and teaching methods employed in the school, but they are in a better position than an outsider would be to coordinate center activities with school practices and to obtain information on pupils from school officials and records.

Needless to say, if competent teachers are to be recruited and retained for supervising activities, they must be compensated fairly for their time and talents. The hourly rate of pay for such supervisors should be better than the prevailing rate for regular classroom teachers in the area.

15. Whenever possible, parents of pupils in the program and other local adults should be used as attendance aides. These non-professional workers can take over many important functions in the program, thus enabling the master teacher to concentrate on the training and supervision of the tutors. In many of our centers, the attendance aide served as a foster parent, encouraging the pupils and helping to create a warm and accepting atmosphere. If funds are available, the aides should be paid for their services, to ensure their continued attendance.

16. The physical setting for the tutorial instruction should be such that continuous observation and supervision are possible.

Tutoring pupils in their homes or in the homes of the tutors is not recommended. Adequate adult supervision and systematic observation of the tutor are impossible in these circumstances. Nor can the safety and well being of pupils and tutors be assured.

Tutorial programs might be established in public or private schools, churches, social agencies, or other public buildings where activities can be centralized and adequately supervised. The activities might take place immediately after school, during early evening hours, on weekends, or during the summer, when schools are not in session. If possible, the rooms used for tutoring should be confined to a single floor of the building.

17. No more than twenty tutors should be assigned to a center, and no more than four tutor-pupil pairs should be assigned to a classroom. In our program the number of tutors per center ranged from 10 to 35. Our experience indicates that a teacher cannot adequately supervise more than 20 tutors at a time. Even if two teachers are employed, as was the case in our largest center, the mechanics of properly allocating materials and teaching aids, keeping attendance records, coordinating tutorial activities with the schedules of classroom teachers, and other administrative details can prove overwhelming when a large number of youngsters are involved.

Although the number of tutors and pupils assigned to a given room will depend largely on the nature of the available physical facilities, our experience indicates that assigning more than four pairs to a classroom is inadvisable: the noise level tends to be excessive, and both tutors and pupils tend to be distracted by the other youngsters in the room.

18. The neighborhood school is a desirable location for a tutorial center provided that efforts are made to counteract the school's negative connotation for the low-achieving pupil. The establishment of tutorial centers in neighborhood schools not only provided Mobilization with classrooms equipped with blackboards and other audio-visual materials but also enabled our supervising teachers and tutors to establish a fruitful relationship with each pupil's regular teacher. Teacher and tutor exchanged valuable information on events in the classroom and in the tutorial sessions. Teachers were asked to make recommendations as to what kinds of assistance should be given by the tutors, and tutors were asked to keep teachers informed as to the progress of their pupils. Many teachers and guidance counselors, furthermore, helped to see that pupils attended the program regularly by reminding them of their appointments.

On the other hand, the school atmosphere may prove a disadvantage to the program in that the low-achieving pupil may regard the tutorial activities as just an extension of his school day. To the degree that he has experienced failure and boredom in school, he may carry over into the tutorial situation his feelings of hostility and resentment.

In anticipation of these problems, Mobilization's tutorial program was presented to pupils in terms of individualized attention and assistance with homework assignments. Before reporting to the centers, each pupil received a postcard from his tutors, saying that the tutor was looking forward to meeting him and helping him with his school work, Early tutorial sessions were focused on creating a warm, accepting atmosphere and building a cooperative relationship with the pupil.

19. The tutorial program should be subjected to empirical evaluation. The present studies have demonstrated the effectiveness of Mobilization's tutorial program. We cannot be absolutely certain that replication of the program in another community will produce similar results, although this seems quite likely. While most organizations will have neither the funds nor the research personnel to engage in lengthy studies, some form of on-going program evaluation is necessary.

Objective evaluation implies a comparison of program participants with a non-participant population. At the very least, the language arts skills of participants should be measured on a before and after basis. Grade-wide norms for the school or city could be used as a standard to evaluate the growth of participants. If data from previous years are available, participants can serve as their own controls and judgments can be made as to whether participation in the program has resulted in an accellerated rate of learning.

Likewise, the growth of participants could be compared with that of a group of non-participants who have been matched on relevant variables. The best evidence of effect, however, will be obtained from controlled experiments which permit the random selection of participants and non-participants.

Clearly, the tutorial situation is a fruitful area for educational research. Further research is needed to determine the long-term effects of the tutorial experience and the value of this approach when applied to other populations.

APPENDIX A

Technical Report on the Pupils

The Design of the Experiment

The study of the pupils was structured as a classical experiment with random assignment of pupils to experiment and control situations. Experimental pupils were provided with tutorial assistance either one or two afternoons a week. Control pupils received no tutorial services. Both groups were tested at the beginning of the program and five months later.

The Samples

The samples in the pupil study were drawn from a population of 2,500 fourth- and fifth-grade pupils who were reading below grade level at the time they were tested. A letter in both Spanish and English was sent to the parents of eligible pupils, explaining the tutorial program and inviting them to apply for their children. The guidance counselor in each elementary school screened the signed applications in an attempt to eliminate pupils with various behavioral problems, pupils of retarded mental development, and pupils with long histories of truancy.

From the remaining eligible population, 410 experimental and 185 control pupils were randomly selected. The parents of each experimental pupil were notified by mail on what day the pupil was to report and to which tutorial center he had been assigned. Parents of control pupils were told that space limitations made it impossible to include their children in the program that year.

As the year progressed, our samples were reduced by 54 experimental subjects and 28 control subjects. The experimental attrition consisted of 16 pupils who failed to report to their centers and 38 pupils who reported but subsequently withdrew from the program. The control attrition consisted of 24 pupils on whom data could not be obtained and four pupils who were assigned to tutors in error as replacements for experimental pupils who had withdrawn from the program. (These were not included in the experimental sample.) Thus, the final samples for analysis included 356 experimental subjects and 157 control subjects. As can be seen from Table 2, the background variables of sex and ethnicity appear to be equitably distributed between the samples. Girls constituted 47 percent of the experimental group and 50 percent of the controls. Approximately 54 percent of the experimentals and 58 percent of the controls were in the fourth grade.

TABLE 2
Sex and Ethnic Characteristics of Experimental and Control Pupils

	Experimental		Control	
	No.	%	No.	%
White - Boys	21	5.9	10	6.4
- Girls	17	4.8	6	3.8
Negro - Boys	56	15.7	31	19.7
- Girls	53	14.9	14	8.9
Puerto Rican - Boys	106	29.8	35	22.3
- Girls	95	26.7	56	35.7
Oriental - Boys	4	1.1	2	1.3
- Girls	4	1.1	3	1.9
Total	356	100.0	157	100.0
Subtotals by Ethnicity				
White	38	10.7	16	10.2
Negro	109	30.6	45	28.6
Puerto Rican	201	56.5	91	58.0
Oriental	8	2.2	5	3.2
Subtotals by Sex				
Boys	187	52.5	78	49.7
Girls	169	47.5	79	50.3

Treatment Contrasts

In certain centers pupils were assigned to treatment situations so as to permit the exploration of two dimensions of the tutorial service. The first dimension concerned the extent of pupil exposure to the tutorial experience. Differences in exposure were obtained by having some pupils attend the program two afternoons a week (four-hour group) and others only one afternoon a week (two-hour group). After attrition, our samples for this portion of the experiment consisted of 100 four-hour pupils, 73 two-hour pupils, and 79 control pupils who came from the same schools as the experimental groups. The median four-hour pupil received 92 hours of tutoring out of a possible maximum of 108 hours. The median two-hour pupil received 46 hours of tutoring out of a possible maximum of 52 hours.

The second dimension that was explored involved a contrast between two strategies of tutorial assistance. In one center special "reading stations" were organized for the conduct of particular phonic and word-attack skill activities. Each tutor was trained to supervise activities at one of these stations. For the first hour of each session the pupil worked at a station to which he had been assigned on the basis of his diagnosed word-attack skill deficiencies.[1] When the pupil demonstrated improvement in one skill, he was directed to another station to work on another skill. During the second hour the pupil worked with his assigned tutor on regular homework and reading activities.

This skill-stations center was paired with another center in which the tutors used the same strategy as was used in all other centers. In this "subject-centered" approach, the tutor used as a teaching vehicle any subject, project, or activity that seemed to interest the pupil. Phonic skill training was provided under both the skill-stations and the subject-centered approach. The essential differences between the strategies were (1) that the skill-stations strategy represented a more systematic emphasis on phonic skill training, and (2) that, while experiencing this approach, pupils tended to move from tutor to tutor.

By randomly assigning four-hour pupils to these two centers and utilizing the 40 control pupils in the feeder schools, we created thre samples to contrast the effects of a skill-stations approach with those of a subject-centered approach and a control situation. Each of the experimental groups initially contained 30 pupils. After attrition, 21 pupils remained in the skill-stations group, 27 in the subject-centered group, and 32 in the control group.

Findings

The expected outcomes of the Homework Helper Program as far as pupils were concerned were classified in terms of changes in achievement, behavior, school marks, and attitudes and aspirations. These data were assigned differential weights in the evaluation. Since the major goal of the program was to improve pupil reading achievement, changes in reading achievement were viewed as the major outcome variable. Because standardized tests were used in assessing reading achievement, greater confidence may be placed in the findings in this area than in the findings on behavioral and attitudinal changes. Indeed, in our view, observed changes in attitudes, aspirations, and behavior are meaningful for purposes of evaluation only if these changes eventuate in increased achievement. These data were therefore regarded as secondary outcomes and were accepted with confidence only if they correlated with reading achievement. As will be shown in the next section, significant improvements in reading skill were observed only in the four-hour sample. Consequently, our analysis of school marks and of behavioral and attitudinal data was restricted to comparisons among the four-hour, the two-hour, and the control samples.

Reading Achievement

Pupil reading achievement was measured on a before-and-after basis, using the New York Tests of Growth in Reading, Level C, Form 1 (revised).[2] This test yields raw scores and grade equivalent scores for reading comprehension, with a grade range from 2.0 to 7.7. The publisher reports a Kuder-Richardson reliability coefficient of .95 for fourth- and fifth-grade pupils and a validity coefficient of .89 with the Metropolitan Elementary Reading Test, Form T, Test 1.[3]

Since little is known about the reliability of standardized reading measures when used with Negro and Puerto Rican youngsters, our first step in the analysis of the reading data was to compute test-retest reliability coefficients by correlating prestudy scores with poststudy scores. As can be seen from Table 3, our estimates or reliability fall some distance below that reported by the test publisher. However, with a five-month period between pretest and post-test (during which pupils were undoubtedly learning at differerent rates), even lower reliability coefficients would be expected. Our coefficients should therefore, be viewed, as minimal estimates of reliability. They are high enough, however, for us to consider the test a reliable measure for samples with a high proportion of Negro and Puerto Rican pupils.

TABLE 3

Reliability Coefficients and Related Data for Specific Samples, New York Tests of Growth in Reading, Level C, Form 1 (Raw Scores)

Sample	Samples Size	Prestudy		Poststudy		r.
		Mean	S.D.	Mean	S.D.	
All subjects	513	22.36	9.01	26.93	10.06	.79
Experimentals	356	22.81	9.41	27.66	10.21	.80
Controls	157	21.35	7.96	25.31	9.55	.80

Table 4 summarizes the New York Growth in Reading Test results for all experimental subjects as compared with all control subjects. Although experimental pupils showed a slightly greater mean change than control pupils, the adjusted difference of 1.03 raw-score points was not statistically significant.

TABLE 4

Comparison of Experimental and Control Pupils, New York Tests of Growth in Reading, Level C, Form 1 (Raw Scores)

Sample	N	Prestudy Mean	Poststudy Mean	Mean Change	Adjusted Mean Change[a]
Experimentals	356	22.81	27.66	+4.85	+4.89
Control	157	21.35	25.31	+3.96	+3.86

F_{obs} (1,505) = 3.22

$F_{.06}$ (1,400) = 3.86

---

[a] Extensive use was made of analysis-of-covariance techniques to adjust change means for differences which might have resulted from the inequitable distribution of contaminating factors between samples. The factors used as covariants were sex, ethnicity, school grade, access to school programs in reading remediation, and prestudy reading level (raw score). In the analysis of covariance, the combined effect of these factors was estimated through a multiple-regression equation. Then, change means were adjusted so as to remove these effects. Using the adjusted change means, an over-all "F" test was computed to determine whether or not significant differences were present. Finally, the Duncan Range statistic was employed to test the significance of the differences between the respective means.

In comparing samples which differed in the extent of exposure to the tutorial situation, we hypothesized that pupils assigned to the program four hours a week would show greater reading improvement than would pupils assigned only two hours a week, and that both these groups would show greater reading improvement than would control pupils. The data for the extent-of-treatment comparisons are shown in Table 5. The difference of 2.74 raw-score points between the change means of the four-hour pupils and the control pupils was significant at the .05 level. The differences of 1.79 raw score points between the two-hour pupils and the controls and .95 points between the four-hour and two-hour pupils were not significant.

The reader is reminded that the over-all comparison of experimental and control samples (Table 4) revealed no significant differences. However, when extent of treatment is taken into account, we find that pupils who were receiving four hours of tutoring weekly made significant gains as compared with control pupils. We conclude, therefore, that tutorial assistance results in reading improvement provided that the assistance is given as often as four hours a week.

TABLE 5

Comparison by Extent of Treatment, New York Tests of Growth in Reading, Level C, Form 1 (Raw Scores)

Sample	N	Prestudy Mean	Poststudy Mean	Mean Change	Adjusted Mean Change
4-hour	100	21.33	27.53	+6.20	+6.23
2-hour	73	22.89	28.22	+5.33	+5.28
Controls	79	21.98	25.46	+3.48	+3.49

$$F_{obs}\ (2,244) = 3.68$$
$$F_{.05}\ (2,200) = 3.04$$

In terms of grade equivalents, the four-hour pupils showed an average of six months' reading improvement in five months' time; the two-hour pupils showed five months' growth, and control pupils showed 3.5 months' growth.

To determine the degree to which a pupil's reading improvement was a function of sex, ethnicity, school grade, or access to school programs in reading remediation,[4] a series of factorial analyses of variance was computed. In each case divisions by extent of treatment were cast in columns, and divisions by the second factor were cast in rows. In all the analyses, columnar differences proved to be significant. Since none of the row differences and none of the interactions between columns and rows proved to be significant, we conclude that extent of treatment defines reading improvement regardless of the pupil's sex, ethnicity, school grade, or degree of access to school programs in reading remediation.

To explore the factor of initial reading ability, pupils in the extent-of-treatment samples were classified as being in the lowest, middle, or highest third in initial reading ability. The effects of initial reading level were examined in a factorial analysis of variance. As can be seen from Table 6, the reading-change means for the three levels of initial reading ability (rows) did not differ significantly. However, sizable differences did occur among the three groups in the lowest third by prestudy raw score. The effects of this trend are mitigated by the lack of differentiation in the middle and highest thirds. Nevertheless, these data clearly indicate that the high growth of the four-hour pupils is most closely associated with those pupils who were the most severely retarded in reading at the beginning of the study.

TABLE 6

Factorial Analysis of Variance of Samples by Extent of Treatment and Initial Reading Level, New York Tests of Growth in Reading, Level C, Form 1, Raw Score Change

Extent of Treatment

		4 Hour	2 Hour	Control	Row Totals
Prestudy Reading Level	Low	N = 37 $\bar{X}$ = 8.05	N = 26 $\bar{X}$ = 6.31	N = 25 $\bar{X}$ = 3.12	N = 88 $\bar{X}$ = 6.14
	Middle	N = 36 $\bar{X}$ = 5.06	N = 20 $\bar{X}$ = 5.95	N = 26 $\bar{X}$ = 4.08	N = 82 $\bar{X}$ = 4.96
	High	N = 27 $\bar{X}$ = 5.19	N = 27 $\bar{X}$ = 3.93	N = 28 $\bar{X}$ = 3.25	N = 82 $\bar{X}$ = 4.11
Column Totals		N = 100 $\bar{X}$ = 6.20	N = 73 $\bar{X}$ = 5.33	N = 79 $\bar{X}$ = 3.48	N = 252 $\bar{X}$ = 5.10

Variance Table

Source	df	MS	F	Probability
Between groups	8	79.25	1.90	NS
Rows (prestudy level	2	88.00	1.90	NS
Columns (extensity)	2	166.00	3.59	.05
Interaction	4	36.36	.67	NS
Within groups	243	46.51		
Experimental error	247	46.26		

The reading results for the treatment strategy samples are shown in Table 7. Although the skill-stations pupils showed the highest mean change, they did not differ significantly from the subject-centered pupils or the control pupils. To determine whether these groups showed differential change in phonic abilities, the Diagnostic Tests of Word Attack Skills were administered both before and after the study. A comparison of samples on the various tests in the battery revealed no significant differences. Although the battery appears to have construct validity,[5] we have no concrete evidence of the degree to which it accurately and consistently measures phonic abilities. If, however, the reliability of the Diagnostic Tests may be assumed, we would conclude that there was no difference between these strategies in their effect on phonic skills, and that neither strategy produced significant changes in these skills.

TABLE 7

Comparison of Tutorial Strategies, New York
Tests of Growth in Reading,
Level C, Form 1 (Raw Scores)

Sample	N	Prestudy Mean	Poststudy Mean	Mean Change	Adjusted Mean Change
Skill-stations	21	18.48	25.53	+7.05	+6.53
Subject-centered	27	22.19	27.67	+5.48	+5.74
Control	32	21.84	24.34	+2.50	+2.62

F_{obs} (2,75) = 2.88
$F_{.05}$ (2,70) = 3.13

School Marks and School Behavior

To determine the effects of the program on school marks and school behavior, data were collected on teacher-assigned grades in reading, spelling, mathematics, science, and social studies; on behavioral ratings in work and study habits, social behavior, and written and oral participation; and on school attendance. This information was gathered from school records for the 1962-63 (prestudy) and the 1963-64 (poststudy) academic years.

We were particularly interested in examining assigned marks in reading since the program was focused principally on this area. However, the data on reading marks proved to be quite incomplete. In our examination of school records we found that teachers had been constrained by Board of Education policy to assign marks in two quite irreconcilable ways. If a pupil was achieving at or above grade level in reading or spelling, the usual alphabetical marks representing "excellent", "good", "fair", or unsatisfactory" were assigned. If, however, the pupil was achieving below grade level in these areas, numerical grade-level scores representing his functional achievement were entered on the record in place of the usual alphabetical mark. Regrettably, there is no satisfactory way of equating the alphabetical marks with the numerical grade equivalents. As a consequence, our samples were reduced by pupils who were given marks under one system in the prestudy period and under the other system in the poststudy period. Furthermore, there is some evidence to indicate that these grading rules were not uniformly applied. Although 86 percent of the pupils in our samples were reading below grade level at the time of their inclusion in the program, only 53 percent were given functional reading scores in place of grades in either the prestudy or the poststudy period.

To facilitate the analysis of pupil school marks, the alphabetical categories were converted into numerical codes ranging from 1 for "unsatisfactory" to 4 for "excellent." Taking these codes as ordinal scores, we computed the degree of change for each pupil by subtracting his prestudy score from his poststudy score. The resulting change scores ranged from -3 to 3, with zero indicating no change. For all the marks and behavior ratings, there was extensive clustering of pupils in the samples at the no-change point in the scale. Differences between samples were tested using Chi Square. None of the differences was large enough to be significant. Also, no differences were found between samples in the number of times absent from school or in the number of times tardy to school.

Attitudes and Aspirations

Before-and-after data on pupil attitudes and aspirations were gathered through the administration of a 44-item "Study of Boys and Girls" questionnaire, adapted from a similar instrument used with junior high school pupils in the Detroit Great Cities School Improvement Project. Since many of the pupils could not read well enough to take the questionnaire by themselves, the entire instrument was read to all subjects.

Each item on the questionnaire was scored by assigning codes or weights to responses, from 1 (most negative) to 3 or 4 (most positive). Change scores were obtained by subtracting the prestudy score from the poststudy score for each item. The distribution of change scores over the treatment extensity samples was examined item by item using Chi Square. None of the differences between the samples was significant.

Since the questionnaire was intended as a measure of change in specific attitudinal and aspirational areas, an attempt was made to cluster items into scales. This was done by summing scores over items that were similar in content. The scales developed were educational aspirations, vocational aspirations, attitudes toward the teacher, attitudes toward school, and attitudes toward socially disapproved acts. For each scale, sample change means were computed.

In our analysis of these data, the first hypothesis was that change in attitudes and aspirations would be a function of treatment extensity; that is, pupils receiving the most treatment would show the most positive change. The data do not support this hypothesis, for none of the differences between sample change means was significant.[6/]

Our second hypothesis was that pupils showing high gains in reading would show positive change in attitudes and aspirations. To test this hypothesis, experimental pupils were classified as "improvers" or "non-improvers" depending upon whether their reading-change scores were greater than that of the average control pupil.[7/] None of the differences was found to be significant.

Our third hypothesis was that particular attitudes and aspirations would serve as preconditions for reading improvement. To test this hypothesis a separate analysis of the prestudy scores was conducted for each of the scales. Here again, no significant differences were found between improvers and non-improvers.

If we assume the validity of our instrumentation, we would conclude that the tutorial program had no measurable effect on pupil attitudes and aspirations, and that none of the attitudes and aspirations measured served as preconditions for reading improvement.

The Tutor-Pupil Relationship

The effect of the tutor-pupil relationship on pupil achievement was estimated in two ways. The first estimate was based on the hypothesis that pupils matched to their tutors by sex and ethnicity would come to view their tutors as "role-models" of academic success and, through identification with the tutors, would be motivated toward increased achievement. In our original assignment of pupils to tutors, an attempt was made to generate four groups: (a) pupils who were matched to their tutors by sex and ethnicity, (b) pupils who were matched by sex alone, (c) pupils who did not match their tutors in sex or ethnicity, and (d) control pupils. To test the hypothesis, the reading-change means for these groups were compared in a 2-by-4 factorial analysis of variance (Table 8). Differences due to the type of matching (rows) were significant, whereas differences due to ethnicity (columns) were not. Interaction between the type of matching and ethnicity was high, but not high enough to be significant at the .05 level of probability. When Table 8 was entered to make comparisons between all pairs of means, only the extreme means of 9.31 for sex-ethnically matched Negroes and 2.84 for Negro controls differed significantly. We conclude, therefore, that sex-ethnic matching makes a difference only for Negroes, who tend to show high gains in reading when matched to their tutors in both sex and ethnicity.

TABLE 8

Analysis of Variance by Ethnicity and Categories of Tutor-Pupil Match, New York Tests of Growth in Reading, Level C, Form 1 (Raw Scores)

Tutor Pupil Matching

	Matched by Sex & Ethnicity	Matched by Sex Only	Not Matched	Controls	Row Totals
Negroes	N=19 $\bar{X}$=9.31	N=20 $\bar{X}$=4.35	N=15 $\bar{X}$=5.93	N=25 $\bar{X}$=2.84	N=79 $\bar{X}$=5.37
Puerto Ricans	N=18 $\bar{X}$=3.89	N=62 $\bar{X}$=5.69	N=24 $\bar{X}$=4.67	N=45 $\bar{X}$=3.13	N=149 $\bar{X}$=4.54
Column Totals	N=37 $\bar{X}$=6.68	N=82 $\bar{X}$=5.37	N=39 $\bar{X}$=5.16	N=70 $\bar{X}$=3.03	N=228 $\bar{X}$=4.82

Variance Table

Source	df	MS	F	Probability
Between groups	7	99.51		
Rows	3	126.94	2.78	.05
Columns	1	35.58	.78	NS
Interaction	3	93.38	2.07	NS
Within groups	220	45.07		
Experimental error	223	45.72		

The second estimate of the effect of the tutor-pupil relationship focused on the attitudes of pupils toward their tutors. During the final week of the program, the supervising teachers administered to the experimental pupils in each tutorial center a research schedule containing 42 questions about the pupil's attitudes and feelings toward his tutor. The schedule was constructed so as to permit the expression of negative as well as positive feelings and attitudes. The responses of pupils in the extent-of-treatment samples were used to test the hypothesis that a relationship exists between the attitudes pupils have toward their tutors and the degree of change that occurs in the pupils' reading achievement. As in the examination of general pupil attitudes and aspirations the responses of pupils in the four-hour and two-hour samples were pooled and dichotomized on the basis of degree of change in reading achievement. For each item in the schedule, Chi-Square comparisons were made between the response patterns of high- and low-change pupils. None of the differences in response patterns was found to be significant. In addition, an attempt was made to scale items in the schedule by clustering items of similar content. Again, high-gain pupils did not differ from low-gain pupils. We conclude, therefore, that pupil achievement is not significantly related to or affected by the attitudes pupils have toward their tutors.

Summary

The major findings of this study may be summarized as follows:

(a) Pupils who were tutored four hours a week showed significant gains in reading as compared with pupils who were not tutored. This effect was most pronounced for those pupils who initially were the most retarded in reading.

(b) Pupils who were tutored only two hours a week did not show significant gains in reading as compared with pupils who were not tutored.

(c) Sex-ethnic matching of pupils to their tutors tended to maximize reading gains for Negro pupils. This trend did not apply for Puerto Rican pupils.

(d) The effects on pupil reading achievement of a "skill-stations" tutorial strategy did not differ from the effects of a "subject centered" strategy. Neither approach resulted in significantly greater gains in phonic skills than were made by pupils who were not tutored.

(e) The tutorial services did not produce a measurable change in school marks, school behavior ratings, or pupil attitudes and aspirations.

(f) The degree of pupil reading improvement was not related to the pupil's measured attitudes toward his tutor.

NOTES

1. The instrument used was the Diagnostic Tests of Word Attack Skills (S. A. Cohen and Robert D. Cloward), unpublished experimental edition, Mobilization for Youth, Inc., 1963. This battery assesses skills in the areas of recognizing initial and final consonants and consonant blends, discriminating among vowel sounds, syllabication, visual memory for short words, and structural analysis of word endings.

2. Bureau of Educational Research, Board of Education of the City of New York, 1959.

3. Manual of Directions, New York Tests of Growth in Reading, Level C (revised), Forms 1 and 2, Bureau of Educational Research, Board of Education of the City of New York, 1959, p. 12.

4. Some of the experimental and control pupils were receiving remedial instruction in school during the period covered by this study. Since our samples do not constitute a random sample of pupils receiving remedial reading in the schools, our findings should not be construed as an evaluation of the remedial-reading programs offered by the schools.

5. By "construct validity" we mean that the content of the test items focuses specifically on particular phonic skills.

6. Since the data appeared to be normally distributed, analysis of covariance was used to correct for sample differences in the prestudy measure.

7. The reading change mean for controls was 3.5 raw-score points.

APPENDIX B

Technical Report on the Tutors

The Design of The Experiment

The study of the tutors was structured as a classical experiment with random assignment of eligible tutor applicants to experimental and control groups. Experimental subjects served as tutors; controls did not. Both groups were tested at the beginning of the program and seven months later.

Sample

The samples were drawn from a population of tenth- and eleventh-grade students who had applied for jobs as tutors. Applicants were deemed eligible for the experiment if (1) they were 16 years old and could obtain working papers, (2) they were in no immediate danger of failing in their own school work, and (3) their school records indicated achievement in reading no more than three years below grade level. The applications were routed through the high school guidance counselor in an attempt to eliminate applicants who were emotionally distrubed or in serious academic jeopardy.

Of the 227 sutdents who met the eligibility requirements, 155 were randomly selected as experimental subjects and hired as tutors. The remaining 72 students, who were placed in the control group, were told that they could not be hired as tutors that year but that they would be offered tutoring positions the following year provided that they participated in our post-study testing program.[1/]

TABLE 9
Sex and Ethnic Characteristics of Experimental Tutors and Control Subjects

	Experimental		Control	
	No.	%	No.	%
White - Boys	34	22.0	18	25.0
- Girls	67	43.2	29	40.2
Negro - Boys	9	5.8	5	7.0
- Girls	14	9.0	8	11.1
Puerto Rican - Boys	12	7.8	5	7.0
- Girls	16	10.3	7	9.7
Orientals - Boys	0	0.0	0	0.0
- Girls	3	1.9	0	0.0
Total	155	100.0	72	100.0
Subtotals by Ethnicity				
White	101	65.2	47	65.2
Negro	23	14.8	13	18.1
Puerto Rican	28	18.1	12	16.7
Oriental	3	1.9	0	0.0
Subtotals by Sex				
Boys	55	35.6	28	39.0
Girls	100	64.4	44	61.0

Instrumentation

To estimate the effects of the program on the tutors, before-and-after-program data were collected on reading skills, school achievement, attitudes, and aspirations. The school-achievement data were obtained from school records; marks for the 1962-63 academic year were defined as prestudy, and marks for the 1963-64 academic year were defined as poststudy. Data on attitudes and aspirations were obtained through the prestudy and poststudy administration of a 54 item questionnaire developed by the present writer.

The Advanced Level of the Iowa Silent Reading Tests (Revised New Edition)[2/] was used to measure changes in the reading skills of experimental and control subjects. This battery measures skills in the following areas: (1) rate of reading; (2) comprehension of words, sentences, paragraphs, poetry, and short articles; (3) rapid reading for specific information; and (4) skills needed for effective use of an index. Raw scores for each of the eight subtests can be converted by tables into standard scores so as to permit meaningful comparison of scores from one subtest to another. The median of the subtest scores is used as a total score for the battery. The publishers report Kuder-Richardson reliability coefficients ranging from .68 to .87 for individual subtests and a total-score reliability coefficient of .93 for tenth- and eleventh-grade students.[3/]

In addition, an experimental edition of the Quick Word Test (Borgatta-Corsini)[4/] was administered on a prestudy basis to measure the verbal aptitude of experimental and control subjects. Although it is relatively new, this short test of verbal facility has been well standardized and validated. The publishers report a Kuder-Richardson reliability coefficient of .92 for tenth grade students, a test-retest coefficient of .88, and correlations ranging from .72 to .85 with other standardized intelligence tests.[5/]

FINDINGS

Attrition

The present study reports on 97 experimental and 57 control subjects for whom reasonably complete data were available. Most of the attrition in the experimental subjects resulted from mid-year class changes which made it impossible for students to attend late-afternoon tutoring sessions. Of the 36 experimental subjects who resigned before the end of the program, 24 did so because they were scheduled for late-afternoon classes or laboratories. Table 10 analyzes the attrition of both samples.

The sex and ethnic characteristics of the samples available for study are presented in Table 11. Comparison of these data with Table 9 indicates a slightly lower proportion of white subjects in the experimental sample and a lower proportion of Puerto Rican subjects in the control group. Since relatively few Negroes were lost, Negroes show a proportional increase in both samples.

TABLE 10

Sample Attrition

Experimenta Subjects	No.	%
Original sample	155	100.0
Failed to report for tutoring job	19	12.3
Resigned before end of program	36	23.2
Incomplete data on standardized tests	3	1.9
Total attrition	58	37.4
Sample available for study	97	62.6
Control Subjects		
Original sample	72	100.0
Did not report for poststudy testing	13	18.0
Treatment confounded (accidently hired as tutors)	2	2.8
Total attrition	15	20.8
Sample available for study	57	69.2

TABLE 11

Sex and Ethnic Characteristics
of Samples Available for Study

	Experimental		Control	
	No.	%	No.	%
White	58	59.8	39	68.4
Negro	21	21.6	11	19.3
Puerto Rican	18	18.6	7	12.3
Boys	31	32.0	22	38.6
Girls	66	68.0	35	61.4
Total	97	100.0	57	100.0

Reading Achievement

The prestudy reading results for experimental and control subjects are presented in Table 12, together with the results on the Quick Word Test. As can be seen from this table, the samples available for study were quite comparable in specific reading skills as measured by the Iowa Silent Reading Tests and in verbal aptitude as measured by the Quick Word Test. Translating the reading-scale score means into grade-equivalent scores, the experimental subjects had a prestudy mean reading level of grade 9.9 as compared with a mean level of grade 10.1

for the controls. Since the mean grade of origin for both samples was 10.7, the average experimental subject was reading eight months below grade level and the average control subject was reading six months below grade level. However, both samples show considerable range in reading achievement. Twenty percent of the experimentals and 21 percent of the controls were reading below the eigth grade level.[6/] In addition, the Quick Word Test means for both samples fall approximately one stainine below the median for tenth - and eleventh grade students in the standardization population. Thus, both samples show slightly lower than normal reading achievement and verbal aptitude.

The degree of reading change for each subject was obtained by subtracting his prestudy standard score from his poststudy standard score on each subtest. Analysis of covariance was used to correct the change means for prestudy differences between samples in reading level and verbal aptitude. As can be seen from Table 13, both samples showed high growth in total scores as well as substantial growth in most of the subtests. Experimentals show significantly more improvement than the controls on three subtests concerned with reading comprehension and study skills. The total score difference is significant well beyond the .001 level of probability.

TABLE 12

Prestudy Means and Standard Deviations for Experimental and Control Subjects, Iowa Silent Reading Tests (standard scores) and Quick Word Test (raw scores).

	Experimentals (N=97)		Controls (N=57)	
Iowa Silent Reading	Mean	S.D.	Mean	S.D.
Rate	176.2	19.31	179.0	18.57
Comprehension	163.0	19.70	163.0	19.10
Directed reading	147.5	18.43	147.4	17.76
Poetry	159.9	19.56	161.6	20.25
Word meaning	143.8	22.09	143.2	18.51
Sentences	174.2	15.71	175.4	14.17
Paragraphs	161.5	18.34	161.3	18.58
Use of index	161.3	17.86	161.8	21.19
Key words	163.9	18.16	163.5	20.29
Total (Median Score)	161.7	14.28	162.7	14.86
Quick Word Test	56.9	14.32	56.0	14.26

When the total standard-score means are translated into grade equivalencies, in the seven months between pre- and post-study administrations, experimental subjects show a mean growth of 3.4 years as compared with 1.7 years for the control subjects.

TABLE 13

Analysis of Covariance of Change from Prestudy to Poststudy, Iowa Silent Reading Tests (standard scores), for Experimental and Control Subjects

	Adjusted Mean Change		Mean Squares		
Test	Experimentals	Controls	Treatment	Error	F
Rate	-4.9	-6.4	83.69	716.40	0.12
Comprehension	9.0	1.4	2058.23	265.10	7.76**
Directed reading	24.8	15.0	3440.00	373.71	9.20**
Poetry	12.0	7.9	611.36	237.99	2.57
Word meaning	12.6	11.4	53.93	230.18	0.23
Sentences	9.9	10.3	5.45	103.67	0.05
Paragraphs	15.3	10.1	963.15	193.42	4.98*
Use of index	18.7	13.8	834.28	227.31	3.67
Key words	13.3	9.6	485.38	203.03	2.39
Total (Median Score)	13.2	7.6	1110.79	78.57	14.14***

*Significant .05
**Significant .01
***Significant .001

Since an alternate form of the same test was used in the poststudy administration, it may be that a substantial portion of the increase for both groups was due to their increased familiarity with the complex directions for taking the test. It should be noted, however, that experimental subjects showed twice as much improvement as the controls. Thus, if we take the control mean as an estimate of the effects of normal growth plus increased familiarity with the test situation, the net effect attributable to the experimental treatment is a reading increment of 1.7 years.

To determine the degree to which reading improvement was a function of initial reading skill, subjects were classified by prestudy reading level (above or below the median) and by treatment (experimental or control). Since prestudy reading levels were highly correlated with ethnicity, splitting at the median for all subjects would have placed most of the white tutors in the "high" group and most of the Negroes and Puerto Rican tutors in the "low" group. To remove the effects of ethnicity, the median splits in Table 14 were made at the median score for each ethnic group, thereby guaranteeing equal proportions of Negroes, White, and Puerto Rican youngsters in the high and low groups.

Change in reading was examined in a 2 X 2 factorial analysis of variance with treatments in columns and initial reading levels in rows (Table 14). Not only were both the row and the column differences significant, but each cell mean differed significantly from all other cell means. Since the interaction of rows and columns was not significant, the differences among cells means may be attributed to the additive effects of prestudy reading level and type of treatment. The data clearly indicate that the experimental treatment was effective for all subjects and that the effects of the experimental treatment were maximized for subjects with initially low reading skill.7/

To examine the effects of verbal attitude, grade of origin, sex, and ethnicity, each factor in turn was used in a 2 X 2 factorial analysis of variance. In all these analyses, treatment differences were significant beyond the .001 level of probability. None of the differences between levels of the respective factors and none of the interactions between rows and columns was significant. Thus, except for the factor of initial reading level, reading change was not influenced to any significant degree by the demographic and intellectual characteristics of the subjects.

TABLE 14

Factorial Analysis of Variance of Reading Change by Prestudy Reading Level and Type of Treatment, Iowa Silent Reading Tests (Total Standard Score)

	Experimental	Control	Row Totals
High **Prestudy Reading Level**	N=48 $\bar{X}$= +11.52	N=31 $\bar{X}$= +4.19	N=79 $\bar{X}$= +8.65
Low Prestudy Reading Level	N=49 $\bar{X}$= +15.37	N=26 $\bar{X}$= +10.69	N=75 $\bar{X}$= +13.75
Column Totals	N=97 $\bar{X}$= +13.46	N=57 $\bar{X}$= +7.16	N=154 $\bar{X}$= +11.13

Variance Table

Source	df	MS	F	Probability
Between groups	3	825.92		
Rows	1	1001.25	10.70	.01
Columns	1	1428.25	15.26	.001
Interaction	1	68.25	.73	N.S.
Within groups	150	93.74		
Experimental error	151	93.57		

School Achievement

Pre- and poststudy academic averages for each subject were computed by summing the subject's grades in academic, commercial, and vocational courses for the 1962-63 and the 1963-64 academic years and dividing by the number of courses taken each year. Prestudy and poststudy mean averages for the experimental and control samples are presented in Table 15. Although both samples showed slightly lower poststudy mean averages, the poststudy means did not differ significantly even when prestudy differences were removed by covariance.

Since the experimental subjects showed dramatic improvement in reading skills, the experimental and control samples were divided at the median reading gain for controls (+7 standard-score points) and their school marks were examined in a 2 X 2 factorial analysis of variance (Table 16). Since no significant differences were found between high and low reading improvers, or between experimental and control groups, we conclude that the program had no effect on school marks even when reading improvement is taken into consideration.

TABLE 15
Analysis of Covariance
of School Marks for 77 Experimental
and 42 Control Subjects*

Sample	Prestudy Mean	Poststudy Mean	Mean Change	Adjusted Mean Change
Experimentals	81.19	77.84	-3.35	-3.37
Controls	80.75	78.55	-2.20	-2.17

F_{obs} (1,112) = 0.393 $F_{.05}$ (1,100) = 3.94

*Samples include only those subjects for whom complete data on school marks could be obtained for the 1962-63 academic year.

The reader may find it somewhat disconcerting that improvement in reading was not accompanied by correlative improvements in school achievement. This finding, however, has reference to immediate or concurrent effects, for the poststudy measure of school achievement consisted of school marks earned during the period of treatment. It is possible that the reading gains made by tutors during the program eventually will be translated into increased achievement as they progress through school. It is at least clear that serving as a tutor did not have a deleterious effect on the tutor's school achievement.

TABLE 16

Factorial Analysis of Variance of School Marks by Degree of Reading Improvement

		Experimentals	**Controls**	**Row Totals**
Reading Gain	High	N=55 $\bar{X}$=77.89	N=23 $\bar{X}$=76.27	N=78 $\bar{X}$=77.41
	Low	N=22 $\bar{X}$=77.72	N=19 $\bar{X}$=81.30	N=41 $\bar{X}$=79.38
Column Totals		N=72 $\bar{X}$=77.84	N=42 $\bar{X}$=78.55	N=19 $\bar{X}$=78.09

Variance Table

Source	df	MS	F	Probability
Between groups	3			
Rows	1	104.19	1.25	N.S.
Columns	1	13.60	.16	N.S.
Interaction	1	172.74	2.09	N.S.
Within groups	115	82.77		
Experimental error	116	83.54		

Attitudes and Aspirations

Changes in attitudes and aspirations were measured through the prestudy and poststudy administration of a questionnaire containing 54 items focused on attitudes toward school and school-related activities, educational and vocational aspirations, and social values.[8/]

The prestudy responses of experimental and control subjects were asymetrically distributed toward positive response categories for most of the items in the questionnaire. For example, when asked how they felt about school, 54 percent of the total sample indicated that they liked it very much and 44 percent indicated that they liked it more than they disliked it. Only 2 percent indicated that they disliked school. Likewise, 90 percent of the experimental and control subjects thought that their teachers were doing a good job of teaching and 72 percent thought that their teachers were really interested in students. Only 10 percent thought of their school as a "dreary place" and 5 percent thought that their subjects were boring.

Both samples showed very high educational aspirations on the prestudy measure. Seventy four percent indicated that they wanted to go to a four-year college and 13 percent aspired to a two-year college. In response to the question of how far they thought they actually would go in their schooling, 67 percent said four or more years of college and 16 percent said two years of college. All of the respondents indicated that they expected to graduate from high school.

In analyzing differences between experimental and control samples, four procedures were followed. First, the responses of experimental and control subjects were compared for each item on a prestudy and a poststudy basis. Experimental subjects did not differ significantly from control subjects in prestudy or poststudy responses to any of the items concerned with attitudes, aspirations, and values, nor did they differ in the degree of change from prestudy to poststudy measures on these items. Furthermore, experimental subjects with high reading gain did not differ significantly from experimental subjects with low reading gain.

Second, items were grouped by content into scales concerned with attitudes toward school, educational aspirations, and social values. The prestudy responses of experimental and control samples were subjected to scalogram analysis following the procedure suggested by Edwards.9/ Two scales were found to meet the criteria of unidimensionality.10/ The first consisted of four items concerned with educational aspirations and the second consisted of three items concerned with socially deviant values. However, experimental subjects did not differ significantly from control subjects in their prestudy or poststudy distributions of scale scores for both scales (Tables 17 and 18).

TABLE 17
Distribution of Aspiration Scale Scores* by Sample

Prestudy Scale Score	**Experimental**	**Control**
4	46 (48.9)	26 (47.3)
3	21 (22.3)	10 (18.2)
2	8 (8.5)	4 (7.5)
1	7 (7.5)	8 (14.5)
0	12 (12.8)	7 (12.7)
Total	94 (100.0)	55 (100.0)

Chi Sq. =2.11 Probability: Not significant

Poststudy Scale Score		
4	47 (50.0)	23 (41.8)
3	15 (16.0)	13 (23.6)
2	11 (11.7)	5 (9.1)
1	13 (13.8)	6 (10.9)
0	8 (8.5)	8 (14.6)
Total	94 (100.0)	55 (100.0)

Chi Sq. = 3.12 Probability: Not significant

---

* The scale consisted of the following items: (1) How far would you like to go in school? (2) How far do you think you actually will go in school? (3) How far do your parents want you to go in school? (4) How far did an adult whom you admire go in school? Responses indicating less than four years of college were considered negative.

TABLE 18
Distribution of Deviant-Value Scale Scores* by Sample

Prestudy Scale Score	Experimental	Control
3	23 (24.5)	18 (32.7)
2	43 (45.7)	19 (34.5)
1	22 (23.4)	14 (25.5)
0	6 (6.4)	4 (7.3)
Total	94 (100.0)	55 (100.0)

Chi Sq. = 2.01 Probability: Not significant

Poststudy Scale Score	Experimental	Control
3	15 (16.0)	14 (25.4)
2	39 (41.5)	21 (38.2)
1	30 (31.9)	17 (30.9)
0	10 (10.6)	3 (5.5)
Total	94 (100.0)	55 (100.0)

Chi Sq. = 2.78 Probability: Not significant

---

* Respondents were asked how much their friends admired three abilities: (1) ability to get good grades in school, (2) ability to keep one's mouth shut to the police, teachers, and others in authority (reflected), and (3) ability to be hard or tough (reflected). The response categories were "not admired" (negative), "don't care" (negative), and "admired" (positive).

Third, a separate analysis of two questions concerned with educational aspirations and expectations was conducted. For each subject, a discrepancy score was computed by comparing his response to the question of how far he would like to go in school (aspiration) with his response to the question of how far he thought he actually would go in school (expectancy). As can be seen from Table 19, the mean discrepancy score for both samples was very small. Experimental subjects did not differ from control subjects in prestudy or poststudy mean discrepancy.

In the present study, the lack of positive findings in the areas of attitudes, aspirations, and social values probably was a function of the high positive attitudes and levels of aspiration exhibited by the samples on the prestudy measure. Indeed, the study suggests that only youngsters who have high aspirations and positive attitudes and values apply for jobs as tutors in the first place. Under these circumstances, we would not expect the tutorial experience to have much impact on attitudes, aspirations, and values.

TABLE 19
Comparison of Mean Responses to the Aspiration and Expectancy Items, 94 Experimental and 55 Control Subjects*

	Experimentals	Controls	Difference	t	Probability
Prestudy					
Aspirations	4.87	4.96	.09	.45	NS
Expectancy	4.68	4.67	.01	.02	NS
Discrepancy	.19	.29	.10	.53	NS
Poststudy					
Aspirations	4.82	4.82	.00	.00	NS
Expectancy	4.53	4.56	.03	.16	NS
Discrepancy	.29	.26	.03	.71	NS

---

* The weights for the response categories were as follows:

1 = Get out of school as soon as possible

2 = Finish high school

3 = Go to a vocational or technical school after high school

4 = Go to a junior college (two year)

5 = Go to a four-year college

6 = Go beyond college to graduate school

Finally, to determine whether the tutorial experience encouraged youngsters to seek careers as teachers, subjects were asked how they felt about teaching as a career before and after the programs. As can be seen from Table 20, 90 percent of the experimentals and 76 percent of the controls displayed some prestudy interest in teaching, but only 29 percent of the experimentals and 24 percent of the controls were actually planning to become teachers. None of the prestudy or poststudy differences between the samples was statistically significant. We conclude, therefore, that the program did not significantly affect the tutor's interest in becoming a teacher.

TABLE 20
"How do you feel about teaching as a career?"

Prestudy

Response Categories	Experimentals	Controls
I think I would not like it. I am not interested in it as a career.	9 (9.6)	13 (23.6)
I think I might like it but I don't think I would succeed.	12 (12.8)	7 (12.8)
I think I would like it, but there are other things I would rather do.	46 (48.9)	22 (40.0)
I think I would like it, and I am planning to become a teacher.	27 (28.7)	13 (23.6)
Total	94 (100.0)	55 (100.0)

Chi Sq. = 5.59 Probability: Not significant

Poststudy

Response Categories	Experimentals	Controls
I think I would not like it. I am not interested in it as a career.	18 (19.1)	12 (21.8)
I think I might like it, but I don't think I would succeed.	9 (9.6)	6 (10.9)
I think I would like it, but there are other things I would rather do.	42 (44.7)	22 (40.0)
I think I would like it and I am planning to become a teacher.	25 (26.6)	15 (27.3)
Total	94 (100.0)	55 (100.0)

Chi Sq. = .37 Probability: Not significant

The Regression Studies

In early discussions about the program, many school officials had advocated the establishment of relatively high academic and intellectual requirements for the tutors, on the assumption that tutors with higher than average academic achievement would be more effective than tutors with lower than average achievement. Since pupils in the four-hour group did show substantial gains in reading, a sample of 38 tutor-pupil pairs [11/] was used to determine whether the pupil gains were correlated with the prestudy characteristics of their tutors. Using a stepwise linear regression, we attempted to predict pupil change in reading from the demographic, intellectual, and attitudinal characteristics of their tutors. The demographic variables consisted of the tutor's sex, ethnicity, and grade in school. The intellectual characteristics included the tutor's prestudy reading level, verbal facility (Quick Word Test score), and prestudy academic average. The attitudinal variables consisted of the prestudy responses to five items in the tutor questionnaire concerned with attitudes toward school, interest in reading, interest in becoming a teacher, and educational aspirations (see Table 21.)

Table 21

Attitudinal and Aspirational Items Used in the Stepwise Linear Regression Equations.

In general, how would you size up the way you feel about school?

1. I dislike it very much
2. I dislike it more than I like it.
3. I like it more than I dislike it.
4. I like it very much.

How far would you like to go in school?

1. Get out as soon as possible.
2. Finish high school.
3. Go to a vocational or technical school after high school.
4. Go to a junior college.
5. Go to a four year college.
6. Go beyond college to graduate school.

How far do you think you will actually go in school?

1. Get out as soon as possible
2. Finish high school.
3. Go to a vocational or technical school after high school.
4. Go to a junior college.
5. Go to a four year college.
6. Go beyond college to graduate school.

Table 21 (continued)

I enjoy reading and do a lot of it.

1. Definitely disagree
2. Probably disagree
3. Probably agree
4. Definitely agree

How do you feel about teaching as a career?

1. I think I would not like it. I am not interested in it as a career.
2. I think I might like it, but I don't think I could succeed as a teacher.
3. I think I would like it, but there are other things I would rather do.
4. I think I would like it, and I am planning to become a teacher.

None of these tutor variables, singly or in combination, predicts pupil change above the level of chance (Table 22). Since the final multiple correlation coefficient of .39 does not differ significantly from zero, we would conclude that pupil reading change was not related to or influenced by the demographic, intellectual, or attitudinal characteristics of their tutors. Although our sample is rather small, the data do not support the adoption of rigid criteria for the selection of tutors.

TABLE 22

Stepwise Linear Regression of Tutor Variables to Predict Degree of Pupil Reading Improvement (N=38)

Step Number	Variable Entered	Regression Sign	Multiple r	df	F	p
1	Interest in Teaching	+	.178	(1,36)	1.18	ns
2	Aspirations (like to)	-	.237	(2,35)	1.04	ns
3	11th Grade	+	.276	(3,34)	.93	ns
4	Ethnicity-Negro	+	.302	(4,33)	.83	ns
5	Aspirations (will go)	+	.326	(5,32)	.76	ns
6	Quick Word	-	.341	(6,31)	.68	ns
7	School Marks	+	.368	(7,30)	.67	ns
8	Prestudy Reading	-	.379	(8,29)	.61	ns
9	Ethnicity-White	+	.385	(9,28)	.54	ns
10	Like to Read	-	.391	(10,27)	.49	ns
11	Sex Girl	-	.392	(11,26)	.43[a]	ns

[a] F-level insufficient for further computation.

Although it is not possible to influence pupil achievement by selecting tutors with certain characteristics, our earlier findings indicated that gains in the tutor's own reading achievement could be maximized by preselecting tutors with initially low reading achievement. To determine whether achievement in the program was associated with other tutor characteristics, these same tutor variables were used in a stepwise linear regression formula to predict the reading change of 71 experimental tutors.[12] As can be seen from table 23, the combined variables predict tutor reading change for above the level of chance. The two best predictors, however, are prestudy reading level and verbal facility, which yield a multiple correlation of .41. The addition of other variables produces only a slight increase in the accuracy of the prediction (r=.52). These data suggest that tutors with low initial reading skill and high verbal aptitude will benefit most from participation in the program.

TABLE 23

Stepwise Linear Regression of Tutor Variables to Predict Degree of Tutor Improvement in Reading (N=71)

Step Number	Variable Entered	Regression Sign	Multiple r	df	F	p
1	Pre-study Reading	-	.225	(1,69)	3.68	ns
2	Quick Word	+	.407	(2,68)	6.74	.01
3	Interest in Teaching	+	.451	(3,67)	5.70	.01
4	Like School	-	.464	(4,66)	4.54	.01
5	Aspirations (will go)	-	.474	(5,65)	3.78	.01
6	Sex Girl	-	.493	(6,64)	3.419	.01
7	Like to Read	+	.503	(7,63)	3.05	.01
8	Ethnicity Negro	+	.509	(8,62)	2.71	.05
9	11th Grade	+	.515	(9,61)	2.45	.05
10	Aspirations (like to)	+	.518	(10,60)	2.21	.05
11	Marks in School	-	.519	(11,59)	1.98[a]	.05

[a] F-level insufficient for further computation.

Summary

The findings on the tutors closely parallel the findings in the study of their pupils. In both studies, the experimental treatment produced changes in reading skill, and the degree of change was partly a function of prestudy reading level. In both studies no differences were found between before-and after-program measures of school marks, attitudes toward school and school-related activities, educational aspirations, or social values, even when the degree of reading improvement was taken into consideration. The reading improvement of the pupils was not related to the demographic, intellectual, or attitudinal characteristics of their tutors. Improvement in tutor reading skills, however, was closely associated with high verbal aptitude and initially low reading achievement.

NOTES

1. Control subjects were paid five dollars for their participation in the testing program.

2. Harcourt, Brace, and World, Inc., 1939. Form Am was administered at the beginning of the program and Form Bm was administered at the end of the program. Lapse time between administrations was seven months.

3. Manual of Directions: Iowa Silent Reading Tests, New York: Harcourt, Brace, and World, Inc., 1943, p.5.

4. Harcourt, Brace, and World, Inc., 1962.

5. Quick Word Test Manual, Harcourt, Brace, and World, Inc., pp. 14-15.

6. Some of the youngsters who met the eligibility criterion of reading no more than three years below grade level when last tested by the schools might not have been deemed eligible had the Iowa Silent Reading Tests been used.

7. When subjects were classified without reference to ethnicity, the row differences in Table 6 were significant at the .05 level of probability.

8. The responses to each item were assigned weights from "1" to "n", with "1" being the most negative or lowest level of response and "n" being equal to the number of response categories for that item. Change scores for each individual were obtained by subtracting his prestudy weighted response from his poststudy response on each item.

9. Allen L. Edwards, Techniques of Attitude Scale Construction, New York; Appleton-Century-Crofts, 1957, pp. 184-187.

10. Coefficient of reproducibility of .85 or better, with at least a 10-percent gain over the minimal marginal reproducibility.

11. The sample necessarily was restricted to experimental tutors on whom complete data was available and who had worked throughout the year with four-hour experimental pupils.

12. The present sample includes the sample of 38 tutors used earlier to predict change in pupil reading skill.